

Weekly Report

May 21, 2017

1 Work

本周主要在修改AMTG的网页程序，由于经过太多次迭代，程序已经比较乱了，未来需要重构一下。目前在AMTG视图中的重排序已经实现，和边的绘制得到了优化。边的绘制优化主要是所有边一样粗细，并且将从上往下的贝塞尔曲线的控制点左移，将从下往上的贝塞尔曲线的控制点右移，避免所有边的控制点都在同一个竖直轴。这样减少边之间的重叠，可以让人方便得追踪一条边的走向，评委也不会误以为我们做了边绑定。

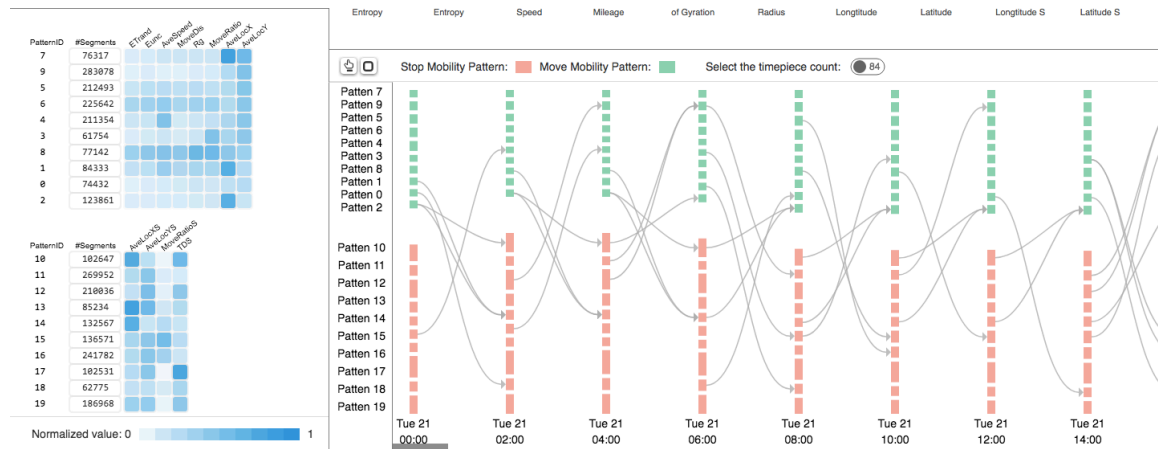


Figure 1: AMTG视图的排序与pixelmap视图保持一致。边之间的重叠已经得到了优化。

| 内容 | 修改方案 | 备注 |
|---|--|--------|
| abstract should be justified | 文字左右两边对齐 | 完成 |
| the discussions are superficial (Section 3.1) | 3.1 Task可能突然出现比较突兀，提供问题的更多细节。 | 简单任务 |
| no design alternatives | 提供一些其他的设计方案。 | 简单任务 |
| conduct a experiment to compare the two methods | 需要对Parallel Set进行基于任务的详细比较。 | 困难任务 |
| what the patterns are | 在Introduction中介绍这个点。 | 完成 |
| related visualization | reviewer提到了两篇文章，可以简单用文字比较一下。 | 完成 |
| regular time intervals | 参考类似的文章，提取重要的时间点，目前打算采用T.von Landesberger论文中的方法。 | 需要继续调研 |
| difficult to follow movement patterns | 需要在代码层面优化一下边的布局。 | 完成 |
| difficult to understand the movements | AMTG图已有图例介绍每个类属于move（绿色）还是stop（蓝色）。 | 完成 |
| an explanation of the ordering algorithm | AMTG图支持重排并且重排后的顺序显示的摆放在AMTG图的左边。 | 完成 |

2 Paper Reading

2.1 How Ideas Flow across Multiple Social Groups

本文使用随机游走方法，在由用户，tweet和单词构成的三元网络中进行游走，计算不同单词之间的转移概率(图2)。a到b的转移概率就是所有从a节点出发的路径中最终到达b的比例。

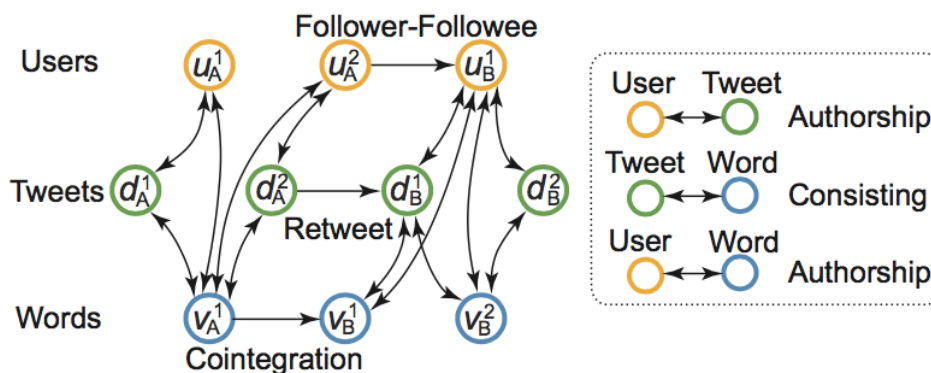


Figure 4: Three-level random-walk-based model for temporal correlation calculation.

Figure 2:

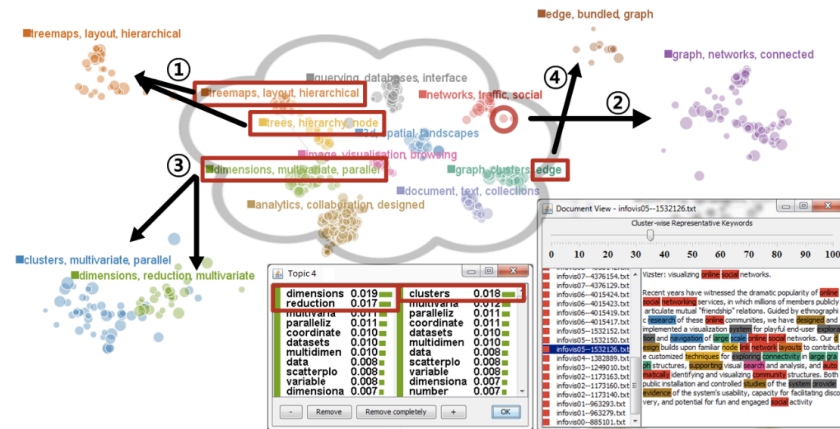
2.2 Weakly supervised nonnegative matrix factorization for user-driven clustering

在本文Introduction的论述中，指出了非负矩阵分解（NMF）方法无论是在结果稳定性还是算法收敛速度上都优于LDA。作者重新定义了NMF的目标函数，将原来无监督的NMF改造成可以接受一定比例数据标记的半监督方法，从而优化结果。

2.3 UTOPIAN: User-Driven Topic Modeling Based on Interactive Non-negative Matrix Factorization

在上一篇文章的基础上，作者将用户交互加入到数据标记结果中，使得NMF在迭代优化目标函数的过程中，对一部分数据人工打上标注，改变目标函数的某些值，是的结果向用户操作靠拢。其中，用户可以进行的操作包括：主题合并，

主题分割，指定文档主题，指定单词主题。我认为将用户交互和机器学习相结合是一个很好的方向，单纯的可视化机器学习算法的结果无法体现出可视分析的重要性。



2.4 MovementSlicer: Better Gantt Charts for Visualizing Behaviors and Meetings in Movement Data

作者使用Gantt Charts展示了用户在不同地点间移动的行为模式。然而Gantt Charts和storyline类似，只能适用于展示少量用户的轨迹（每个用户都需要一条贯穿时间线的曲线）。

